



# **Natural Language Processing (NLP) and Its Impact across Industries – Unlocking the True Potential of Digital Healthcare (A Case Study Approach)**

**Sourajit Roy<sup>1</sup>, Pankaj Pathak<sup>1\*</sup> and S. Nithya<sup>1</sup>**

<sup>1</sup>*Symbiosis Institute of Digital and Telecom Management, Symbiosis International (Deemed University), Pune, Maharashtra, India.*

## **Authors' contributions**

*This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.*

## **Article Information**

DOI: 10.9734/JPRI/2021/v33i35B31906

### **Editor(s):**

- (1) Dr. Paola Angelini, University of Perugia, Italy.
- (2) Dr. Debarshi Kar Mahapatra, Rashtrasant Tukadoji Maharaj Nagpur University, India.
- (3) Dr. Aurora Martínez Romero, Juarez University, Mexico.

### **Reviewers:**

- (1) Shoulin Yin, Shenyang Normal University, China.
- (2) Rustam, S. Si, Telkom University, Indonesia.
- (3) Gulzar Ahmad, Pakistan.

Complete Peer review History: <http://www.sdiarticle4.com/review-history/69545>

**Original Research Article**

**Received 02 May 2021**

**Accepted 06 July 2021**

**Published 07 July 2021**

## **ABSTRACT**

During the advent of the 21st century, technical breakthroughs and developments took place. Natural Language Processing or NLP is one of their promising disciplines that has been increasingly dynamic via groundbreaking findings on most computer networks. Because of the digital revolution the amounts of data generated by M2M communication across devices and platforms such as Amazon Alexa, Apple Siri, Microsoft Cortana, etc. were significantly increased. This causes a great deal of unstructured data to be processed that does not fit in with standard computational models. In addition, the increasing problems of language complexity, data variability and voice ambiguity make implementing models increasingly harder. The current study provides an overview of the potential and breadth of the NLP market and its acceptance in industry-wide, in particular after Covid-19. It also gives a macroscopic picture of progress in natural language processing research, development and implementation.

\*Corresponding author: E-mail: [2pankajpathak@sidtm.edu.in](mailto:2pankajpathak@sidtm.edu.in);

**Keywords:** NLP; sentiment analysis; neural machine translation; pragmatic analysis.

## 1. INTRODUCTION

Natural Language processing has shown great potential in the 21<sup>st</sup> Century enabling newer avenues of A.I. (Artificial Intelligence) and Deep-learning algorithms and has been instrumental in medical patient screening, market intelligence, advertisement, etc. In the last couple of years, we have come across tremendous advancements in the field of NLP, where its applications evolved from just being restricted to the field of research to find their place in real-world business solutions [1]. Natural Language Processing is a subset of linguistics, cybernetics, advanced computing and artificial intelligence that analyses the interactions between machines and human beings and designs a program to compute and analyze huge amounts of natural language data bridging the gap between human language and computer processing. This large unstructured dataset consists of texts, emails, documents, research papers, legal papers, blogs, voice recordings, and social media posts, etc. It is the largest human-created data source that is growing by the minute- exponentially [2].

According to Accenture, Natural Language Processing Markets are expected to grow by \$16 Billion by 2021. The increasing demand for NLP applications is backed by evolving technological advancements ranging from open source frameworks to cloud-based APIs [3]. Its real-life applications cover text extraction from documents, chat bots in customer support services, virtual assistants across computing platforms, intelligent document analysis techniques like OCR (Optical Character Recognition), and sentiment analysis of customer reviews on various social media platforms. Each one of us has at least once interacted with an NLP algorithm through search engine algorithms (Bing, Google, Yahoo!), interacted with a smart assistant (Alexa, Siri, Google), translated one language to the other (Google Translate), or extracted vital information (Health records) from a website [4]. These programs that we have mentioned have shown how a pool of unstructured data can be transformed into exceptional insights. In 2019, a group of scientists from IISc Bangalore, in collaboration with Carnegie Mellon University, proposed a supervised machine learning model to perform extended Word Sense Disambiguation (WSD) over a continuous sense embedding space. A group of researchers came up with a Dialogue

GCN-based emotion recognition engine that could facilitate effective dialogue formulation [5].

A generic NLP modeling approach would start with pre-processing the unstructured data by using Splitting, Reduplication, Normalization, and Stratification and would move on to Parsing the data by using techniques of Tokenization, Misspelling analysis, Parts of speech tagging, etc. The next step involves analyzing the trends by use of Singular voice decomposition, word embedding, and categorization, etc., which leads to the extraction & interpretation of information with the help of chatbots and text-to-speech engines. The final step of the NLP model is Developing Conversational Systems using sentiment analysis, entity recognition, text summarization, etc [6].

## 2. LITERATURE REVIEW

Natural Language Processing or NLP dates back to the 1900s when renowned scientists like Ferdinand de Saussure, Albert Sechehaye, and Charles Bally, Alan Turing made considerable breakthroughs in the form of scientific discoveries and research to formulate the idea of Artificial Intelligence and Natural Language Processing. Though initially, NLP was considered a separate unit from the text Information Retrieval system, which is a highly adaptable statistical-based technique to index and searches huge text samples efficiently. Natural Language Processing (NLP) is a human-computer intercommunication technique involving an extensive range of applications, such as text classification (part-of-speech) tagging, neural machine translation, sentiment analysis, etc [7]

Sentiment Analysis (S.A.) is one of the branches of NLP. It is a growing field of research in context with text-mining techniques wherein the computational analysis of sentiments, notions, and perceptions are determined based on textual data. This textual data can be gathered from various sources like app reviews from mobile computing platforms, various reviews of products on websites, perceptual blogs/posts/tweets on social media platforms, etc [8] Sentiment Analysis can be truly classified based on three levels, namely, aspect-level, document and sentence level. There are two primary approaches adopted for S.A. – the Machine Learning approach and the Lexicon-based

approach. Machine Learning generally involves the analysis of statistical techniques to determine parts-of-speech, entities, sentiments, and other facets of textual analytics. The lexicon-based approach requires a glossary of positive and negative sentiment value assigned, which will help in representing the text [9].

Neural Machine Translation (NMT) is one of the sub-categories of computational linguistics and deep learning, where the fundamental focus is deciphering one language to another [10]. While Google Translate is a dominant industry leader in this field, there are multiple companies who are working in the same space with advanced use-cases being implemented every day. The essential algorithm in every NMT approach is an underlying deep learning application where extensive amounts of datasets are trained to be able to devise a model capable of translating among various languages. One of the more entrenched variants of NMT is the Encoder-Decoder Structure which consists of two recurrent neural networks (RNNs) used together in sync to create an NMT translational model. NMT finds various uses in Network and App-based machine translation based on consumer use, B2B domain-specific machine translation, Real-time text-to-text translation, etc [11].

Though the first known breakthroughs happened in the 1950s with the "Turing Test," until 1990, Natural Language Programming was restricted to NLP and Machine Translation theoretical concepts. It was only after the technological boom of the 2000s, and there was monumental groundwork done on NLP [12]. The growth of NLP can be highly contributed to the establishment of machine learning, artificial intelligence, and steady growth in computational

power, which saw the research and applications of NLP grow manifold. The recent breakthroughs in NLP include: Named Entity Recognition (NER), Parts of Speech (POS) Tagging, Text Classification, etc. To perform NLP tasks in an efficient manner, there is often an overlap between the fields of artificial intelligence, deep learning, and NLP. The ACL Conference (Association for Computational Linguistics) 2019 is a worldwide event that showcases the best research results in the field of NLP, and it is represented in Fig. 1 computational linguistics with 1,544 research papers submissions. "Open Kiwi: An Open Source Framework for Quality Estimation" won the award for best demo paper at ACL 2019 [13].

### 3. METHODOLOGY

#### 3.1 NLP Applications based on Technology

The primary aim of NLP is to simplify human-to-machine interaction as a result of which talking to a machine can be as simple as talking to a human. NLP drives on unstructured data and tries to make sense to a machine. There are a number of use-cases in NLP which prove to justify its market potential. According to IDC, the global data ecosystem would grow to 163 ZB (zettabytes) of data by 2025, which was previously 16.1 ZB in 2016.

We can broadly classify technological applications of NLP into three primary categories: Conversational A.I., Text Analytics (Modelling, Categorization & Clustering), and Named Entity Recognition (NER).

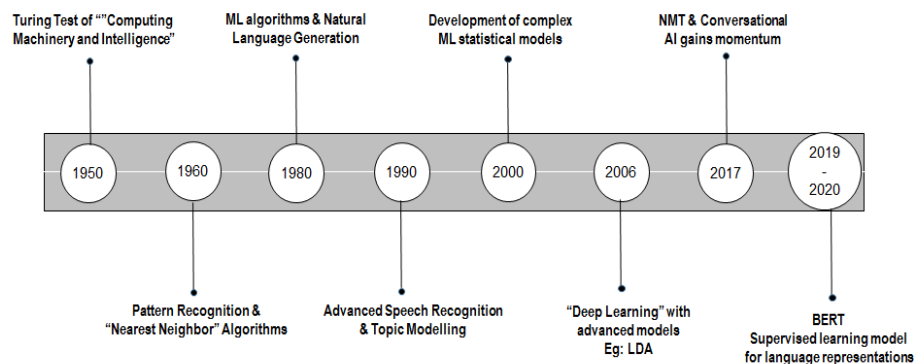


Fig. 1. Evolution of NLP

### **3.2 Conversational Artificial Intelligence (A.I.)**

Conversational systems or A.I. is the consolidation of machine learning, speech recognition, natural language processing to interpret, process, acknowledge text or voice input in logical ways and figure out genuine responses for a user input typically used in conjunction with Intelligent Virtual Agents (ICUs) [14]. Probably the best use-case of Conversational A.I. can be found in multi-channel businesses so that a customer can start a conversation in one channel and carry on with the conversation in another, leading to a better user experience. These multi-channels consist of voice assistants, IVR, SMS, Emails, Web/Mobile hooks, Social messengers, etc.

### **3.3 Text Modelling, Categorization, and Clustering**

Text Modelling refers to the statistical method, primarily Unsupervised Learning, of discovering hidden topics from long textual documents. One of the popular methods of Text Modelling includes Latent Dirichlet Allocation (LDA), which builds models based on topic per document and words per document, modeled as Dirichlet Distributions. This technique was used by the Securities and Exchange Commission (SEC) to identify discrepancies in disclosure reports of companies charged with financial misconduct.

Text Categorization or classification or tagging sorts a specific text into constituent taxonomies after being trained by a human model, which translates to grouping specific pockets Automatic security classification of confidential documents of governments/enterprises is a classic example of Text Categorization [15].

Text Clustering is a similar method to text categorization, where a grouping of texts or documents occurs based on similarities in context. Its uses vary from documents to social media posts to blogs to discover hidden subjects. The Centre for Tobacco Products (CTP) used text modeling, categorization, and clustering to classify and group its documents based on certain keywords.

### **3.4 Named Entity Recognition (NER)**

Named Entity Recognition is an information extraction approach that identifies entities in a

text and groups them into predefined categories. These entities can be names, quantities, people, locations, days, times, monetary values, etc. It is really useful in analyzing texts from unstructured documents and designates them into known categories. E.g., the word "Darwin" can signify someone's name as well as a city in Australia. NER can be used in such cases to classify between the two sets of uses of the word "Darwin." Government agencies worldwide use NER in social media platforms to identify threat perpetrators for terrorist attacks and cybercrimes [16].

## **4. NLP APPLICATIONS BASED ON INDUSTRY**

### **4.1 Automotive**

Self-driving cars were one of the most significant developments in this particular industry. In the past few months, A.I. and NLP have further revolutionized this sector with "in-car assistants," which develop the ability to mimic the human mind in on-road situations over time, allowing the automobile to respond to voice commands and infer on-the-road actions without human intervention. In 2020, autonomous vehicles created some major strides with companies like Waymo, Active, and were expected to offer autonomous taxis. Autonomous vehicles also received a major boost after being featured in Gartner Hype Cycle 2019 [17].

### **4.2 Use-Cases**

Mercedes-Benz launched its MBUX UI interface for its cars which enables a virtual assistant with the help of artificial intelligence and NLP to learn from the driver's habits and present personalized recommendations and offers, taking into account driving behavior, tone of voice, etc. of the driver.

### **4.3 Retail**

The retail industry is currently going through a phenomenal changeover adopting digital transformation measures at all levels. The industry had been an early adopter of NLP through chat bots and communication interfaces. There has been a leap in NLP technological innovations now like text categorization, text clustering, sentiment analysis, inferences, cognitive adaptation, and predictions. NLP has unleashed interactive analytics platforms which allow retailers to take advantage of this solution. With aspects like text analysis, sentimental analysis, NLP-powered chatbots, chat-based

product recommendation, advertising, compliance mandates, etc. [18]. The executives and C-suite can make better-informed decisions with the insights gained, thereby improving product and service offerings to flourish in the market. It serves as both a cost-effective solution and a time-saving process to interact with customers and unstructured data to gain knowledge on trends and drive personalized engagements.

#### 4.4 Use-Cases

Enterprises are adopting several NLP algorithms given the need for adaptation during this pandemic era. Among them, sentiment analysis, otherwise called open mining, has been proposed by several retailers, especially e-commerce/m-commerce retailers, in order to gauge the customer perception of their brand. This gives a broad scope of outputs like aligning their market strategy, sales targets by recognizing the emotions, views, or attitude. It could quantify the sentiment or classify it at a more granular level (e.g., ranging from being very positive to very negative) of the overall customer base. By interpreting the enormous data from everywhere, especially social media, high-level predictions on quantifiable KPIs and revenue levers can be achieved. Besides this, another major application achieved is crisis management. As the brand's timely reaction and perception are measured, any possible crisis could be averted by managing the right decisions at the right time [19]. Leading E-commerce brands like Amazon, Myntra, eBay uses sentiment analysis to monitor the brand image continuously and efficiently manage the ecosystem. Leaders like Amazon use and offer solutions like Amazon extract, which is an NLP-powered search index, and Amazon comprehends sentiment analysis to enhance applications.

#### 4.5 Banking and Finance

The banking ecosystems comprise various verticals like Payment services, Loans, Cash Management, Insurance, etc., encompassing various functions like Cash transactions, Remittance, Payment gateways, Branch Banking, Customer Servicing, etc. In 2020, the field of A.I. and NLP had brought in advanced cognitive technologies, which has impacted and improved all of the above banking functions with Cognitive Computing improving customer

activation through personalized services. Cognitive computing, along with the currently implemented ERP systems, can upend operational finance with remarkable speed, agility, and transparency to finance processes. These technologies find high value in the fields of Internal Audit and Financial Planning & Analysis (FP&A) functions of financial institutions [20].

#### 4.6 Use-Cases

Enterprises are now developing financial systems for designing a credit score model for under banked clients. Lenddo EFL allows financial institutions to offer loans and credit facilities for the middle-class man who has no credit history and helps banks understand these customers' credit risk based on their digital footprints [21]. Lenddo EFL has recently developed FICO (Fair, Isaac, and Company) Score services in India by allowing them access to digital customer data with their agreement. There also have been reports of high developments of sentiment analysis-based trading platforms like the Sigmoidal trading platform that automates mining information stock price predictions based on social media and news reports using document classification and named-entity recognition providing relevant information in accordance with client needs.

#### 4.7 Healthcare

Although NLP has been in use across all sectors and industries, its use in the Healthcare industry has been of utmost importance.

The covid-19 pandemic has seen a paradigm shift in the use of modern technologies and has worked on patients' convenience. In India, the Arogyasetu app was launched for the correct identification and prevention of corona virus cases worldwide. Healthcare provider organizations have implemented online portals with chatbots, virtual health assistants (VHAs) that use NLP, sentiment analysis, and concept analysis to create a personalized experience for them [22]. A significant rise in the adoption of personal health assistants has also been seen where customers are provided ready to use the information on diseases, symptoms, causes, and cures. In some cases, emergency ambulance services and doctor-on-call are also provided. Similar applications of NLP can be found in pharmaceutical sciences wherein the machine can sort through thousands of pages of clinical

research allowing A.I. systems to choose which compounds might be most effective for biological targets and aid in the process of drug delivery. Recently, "machine-vision" image analysis was used to help in finding cures using drug discovery to battle deadly viruses such as Ebola or Zika and can be very beneficial in finding cures for Covid-19 in the coming days. We will now discuss the healthcare aspect of NLP and the specific use-cases of the Healthcare industry of NLP in detail in the next section.

#### 4.8 Digital Healthcare with NLP – The Case Study

The healthcare industry has been facing a number of challenges especially bearing the effects of Covid-19. With the increased cases, the workload on physicians and technicians has been shot up, leading to an awakening call for the industry to avoid time-consuming manual processes which have the scope of automation. It has been forced to drop down the risks of communicable diseases throughout the patient's journey while retaining their experience:

- Requires retrospective analysis beforehand
- Lack of data-driven inputs, with the availability of large unstructured data sources
- A time-consuming process to understand the history of a patient and document it.
- Lack of unified format and unified conversion of variability in language
- Data availability includes the history of patients, the factors influencing a patient's lifestyle, disease terminologies, and lab results related to it, medications recommended, the summary of discharge reports.

- Consumers have shifted their focus more on their convenience with health & hygiene.
- Avoid physician workload/burnout in unwanted manual processes.
- NLP in the Healthcare industries comprises of three primary components: Patient-oriented, Clinical-oriented, and administrative oriented. Here we will be addressing the applications of NLP use cases, especially in the viewpoint of the following segments:
  - Physicians / Lab Technicians
  - Researchers / Academic Scholars
  - Patients

#### 4.9 Use Cases of NLP

The industry has unstructured data which are not used to its full potential. From the view of a doctor, tracking the patient's history, previous medications, and lifestyle determinants will speed up the consulting and recovery process. Many companies are thinking on the lines of implementing NLP and A.I. to a vast array of use cases with the help of NLP approaches like doc2vec, which essentially compares and detects changes in clinical reports, Unified Medical Language System (UMLS) leveraged NER extracting clinical concepts from e-medical records, lab reports and discharge summaries. Moreover, extended stacks of LSTM networks (Long-Short-Term-Memory) can be used to link clinical diagnosis and concepts with codified guidelines and establish human-to-machine NLP instructions for robot-assisted surgeries using deep Q-networks [23] and analysis of NLP in healthcare is shown in Table 1.

**Table 1 . Reason for analysis of NLP in healthcare**

<b>What is to be done? The need of the hour</b>	<b>Why? The true purpose of NLP in Healthcare</b>
Structure Data: Drag and drop NLP queries to import data Documentation process : NLP identifies the undocumented details and adds on.	Reduces the overall administration cost post implementation
Prediction: By inferring the lifestyle determinants of a patient NLP predicts the outcomes	Reduces time from weeks to days per patient
Clinical Decision Support: With the unified source of medical history, medication and necessary data, moving to virtual consultations when necessary is made possible	Predicts outcome of patients with the insights derived from data leading to better quality delivery

## 5. MINING MEDICAL RECORDS: GOOGLE DEEP MIND& IBM WATSON HEALTH

### 5.1 Used By

Physicians, Lab Technicians, Medical Consultants, Researchers, Insurance Agents & Patients Tech giants such as Google and IBM have been addressing the most challenging and complex problems in the health sector. Fig. 2 shows a wider view of how the tech giants have accomplished solutions using NLP. The data availability is huge in every medical institution, and they have large unstructured data, average semi-structured data, and very low structured data. Google Deep mind Health, with some free open source programs, has been collaborating at the national level in the United Kingdom. IBM Watson employed a combination of machine learning and NLP capabilities [24].

The clinical notes are highly unstructured data sources that are not directly analyzable. The first focus is to convert it to structured form (i.e.) unstructured text to understandable electronic medical record (EMR). Through this output, the NLP produces structured data for ML & A.I. techniques as machine-readable forms. The previously missed/incompletely coded records are also identified using NLP and completed.

Secondly, it synthesizes the contents from pages of doctor's prescriptions and notes into reduced important bullet points. This reduces the reviewing time significantly of a doctor/administration from weeks and months to days. This allows the staff members and medical experts to focus on complex work rather than this redundant manual work [25]. The key here is that NLP allows translation and understanding of clinical terms and physician notes into insights. The lab results, medical tests, medical images, x-rays, etc., all contribute to the structures or semi-structured data source. By combining the unstructured data with this, NLP would predict and provide insights on the possible outcomes. The accuracy of this depends on the volume and accuracy of data sources.

Patient-specific information through a unified source of electronic medical records accessible anywhere irrespective of language variations It thereby achieves improved clinical documentation supporting the various clinical decisions making it effective in cost and time.

### 5.2 Clinical Decision Support

#### 5.2.1 Infera – advanced CDS system

Used by Physicians, Medical Consultants, Patients

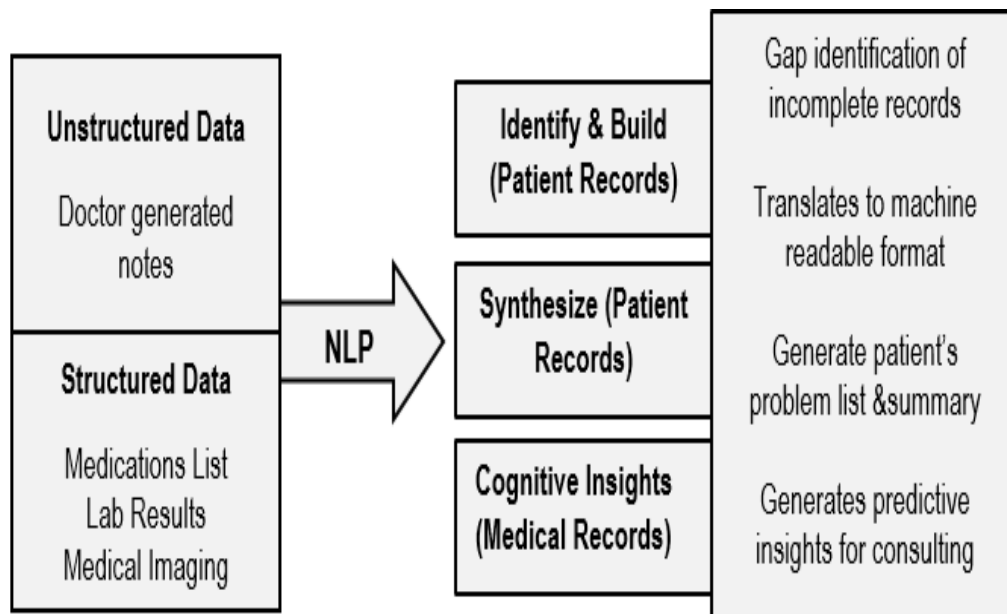


Fig 2. Mining of electronic medical records

Clinical Decision Support is a sub-domain of the broader NLP to assist health professionals by taking over routine tasks through unified format sheets, etc. It evaluates the clinical guidelines of a patient concerning history, allergy, and findings to provide evidence-based diagnosis along with reminders and alerts to assist healthcare providers and patients. This helps to formulate structured data at the entry-level by providing the choice to enter data with controlled uniform vocabularies manually. The patient data includes referral sources, prescribed medications, health professional provided instructions, number of visits and reasons, and progress made, discharge notes [26]. This avoids confusion among other professionals or lab technicians, and pharmacists and improves the quality and efficiency of services, as seen in Fig. 3.

To fully understand the Clinical Decision Support (CDS) system, let us discuss the innovative case of Infera – a U.K.-based CDS system. Infera, by its name, is an advanced CDS system that consistently consolidates all your Electronic Health Record (EHR) Data – both Primary and Secondary into the clinical workflow and provides physicians, medical consultants, and patients with confirmation based clinical guidance and suggestions for medical diagnosis. This is done by combining data that is structured from the Continuity of Care Document (CCD) and evaluating unstructured data found in clinical writings using NLP, which gives the medical professional an exhaustive overview of the

patient's condition [27]. Another highlight of Infera is that it fully integrates which the EHR (Electronic Health Record) of your system to provide fast insights about the patient's health from the pool of unstructured data and facilitates a fully extensive analysis of the patient in the form of an insightful report.

### 5.3 Consultation and Customer Support: Babylon Health

Used by Patients, Administrations of hospitals, Clinics, etc.

In this modern age, patients demand a faster response from their healthcare service providers, and this acts as a key differentiator while choosing between healthcare service providers. Modern-day healthcare applications should be able to apprehend, store and share customer/patient data over the cloud interface, allowing and Healthcare Providers (HCPs) for easy accessibility, which will have the ability to enhance customer care, patient health results. When the patient inputs their specific symptoms and queries, these AI-based apps will evaluate these details over a wide array of government-approved central medical databases, which will contain all kinds of information on various diseases by leveraging advanced NLP capabilities like Speech Recognition, Semantic Analysis, and NER [28].

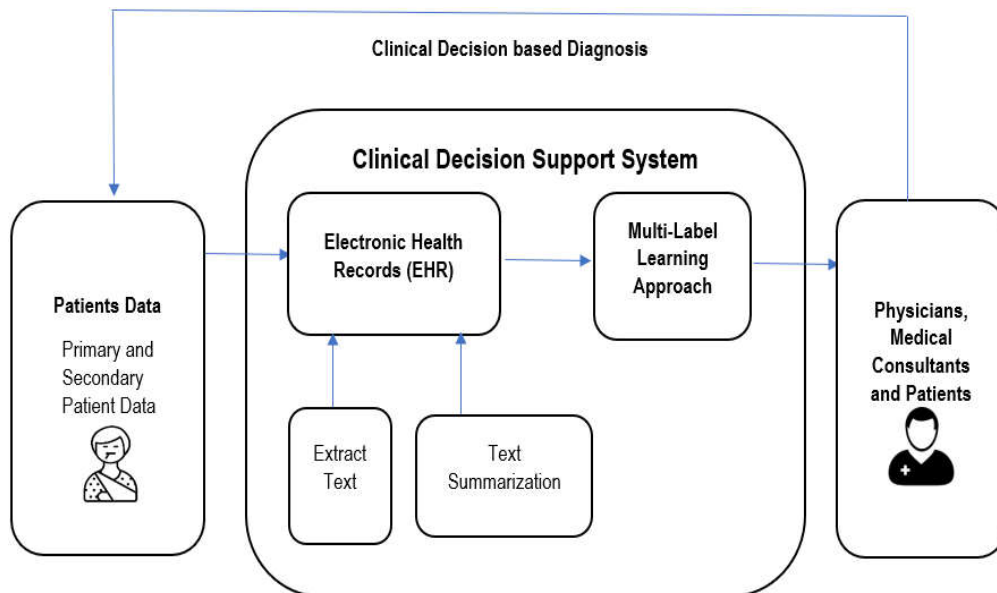


Fig. 3. Clinical decision support (CDS) model pipeline



To understand the consultation and customer support part of NLP, we will discuss Babylon Health. This use case can be implemented across channels like a patient portal, mobile application, etc. The patients can avail of virtual consultations when necessary, and this platform would make it possible by integrating NLP in every step. Companies like Babylon have been launching remote consultations across China, the USA, and Middle East countries. They have bridged the gap between doctors and patients and made the journey easier and convenient. One of the most amazing features of Babylon Health is its flagship Inference Engine which reads user data and simply understands how patients expressed their symptoms, and risk parameters are simply not enough to provide someone the possible matching diseases accurately [29]. Babylon Health's Inference Engine uses an assortment of ML systems that are able to conduct reasoning at the speed of greater than 100 seconds and computes billions of combinations of symptoms, expression of user's symptoms, and take into account the probable risk factors on a per-second basis to help classify condition which may match the information entered providing health information to millions of patients.

As we can see from Fig. 4, the patient journey begins from the moment he/she logs into their Healthcare Provider (HCP) app and inputs patient-specific symptoms/queries. Upon this, the speech understanding and conversational A.I. management system of the app gets enabled and starts with its well-defined processes of Domain Recognition, User Intent Identification, and Dialog State Tracking. At this point, the application provides the patient two varied options, the first being to use the platform-specific A.I.-based patient support system. Upon choosing this option, interactive NLP chat bots are deployed to interpret speech and text and return genuine health responses [30]. The second option one can opt for is to avail the doctor consultations – both online and offline modes are available. By selecting the desired option, the app starts locating medical professionals, doctors, or consultants based on proximity, condition of emergency, priority, etc., and books appointments for the patient. For extreme situations, the application can also connect to emergency health services like an on-the-go ambulance with 24x7 on-call doctors. This entire process has to happen in a very brief span of time and is very helpful for any kind of medical situation.

## 5.4 Managerial Implications of NLP

In recent times, data is now the most valuable resource for any organization, and it is imperative that it capitalize on it in whatever way possible. Though data has been collected in many forms, the major chunk of data is in textual form ranging from social media posts, online reviews, blogs, scientific publications, process, and product specifications, etc., and just shows how pivotal it is to analyze this textual data. This is where NLP comes into the picture. NLP has gained traction not just in the evolution of its underlying technologies to its applications across various industries, but it has evolved within an enterprise and has been absorbed in improving its various verticals to facilitate its corporate objectives:

## 5.5 Hiring and Recruitment

Natural Language Processing (NLP) can significantly optimize the entire process workflow of Human Resources, one of the major activities being recruiting employees for the organization by accelerating the hiring process by intelligent applicant search and filtering out resumes based on organization-specific keywords so that recruiters only have to scan the relevant resumes that match the job description.

## 5.6 Marketing and Advertising

By evaluating the e-footprint of an organization from social media posts, blogs, news articles, keywords, and browser behavior, NLP can be highly beneficial in identifying the most effective digital mediums for maximizing a company's outreach and scoping in newer audiences to showcase their products. Furthermore, it can be used to study the already deployed market campaigns and ads on social media platforms (Twitter, Facebook, Instagram, etc.) to assess its profitability from an organizations point of view

## 5.7 Research and Development

Modern R&D departments can no longer rely on traditional systems of simple digital media monitoring to figure what the modern consumer needs, which are the untapped markets, and what new horizons are opening up in that specific industry and sector. This can be done by semantic analysis and named entity recognition of millions of research papers, whitepapers, news articles, and blogs and generate necessary insights for the R&D team.

## 6. RESULTS AND DISCUSSIONS

### 6.1 Challenges and Future Scope of NLP

The recently conducted Association for Computational Linguistics (ACL) Conference showed great promise of NLP in areas like Fast Domain Adaptation of Quality Assurance (Q.A.) systems using automatic question generation (Q.G.), In-depth insight evaluation using BERT (Bidirectional Encoder Representations from Transformers), Automatic FAQ retrieval and Taxonomy Construction. Although there is a lot of promise shown by NLP, the current adoption rate of NLP systems is relatively low. There are still quite a few bottlenecks that organizations face while implementing NLP models. We know

that most unstructured textual data used in enterprises are in the form of PDF documents. Now, PDF documents are a bit complex and difficult to analyze as they come in many forms, document structures (like titles, sections, headers, and lists), and mostly have embedded visualization like graphs and charts. Moreover, unlike structured data consisting of numbers, NLP data have to work with textual languages, which are tricky in nature as languages differ between countries, with every specific region having a different set of linguistic additions to those specific languages. This is one of the biggest challenges of NLP as they need to support multiple languages without the need to undergo training every time they have to face a new language.

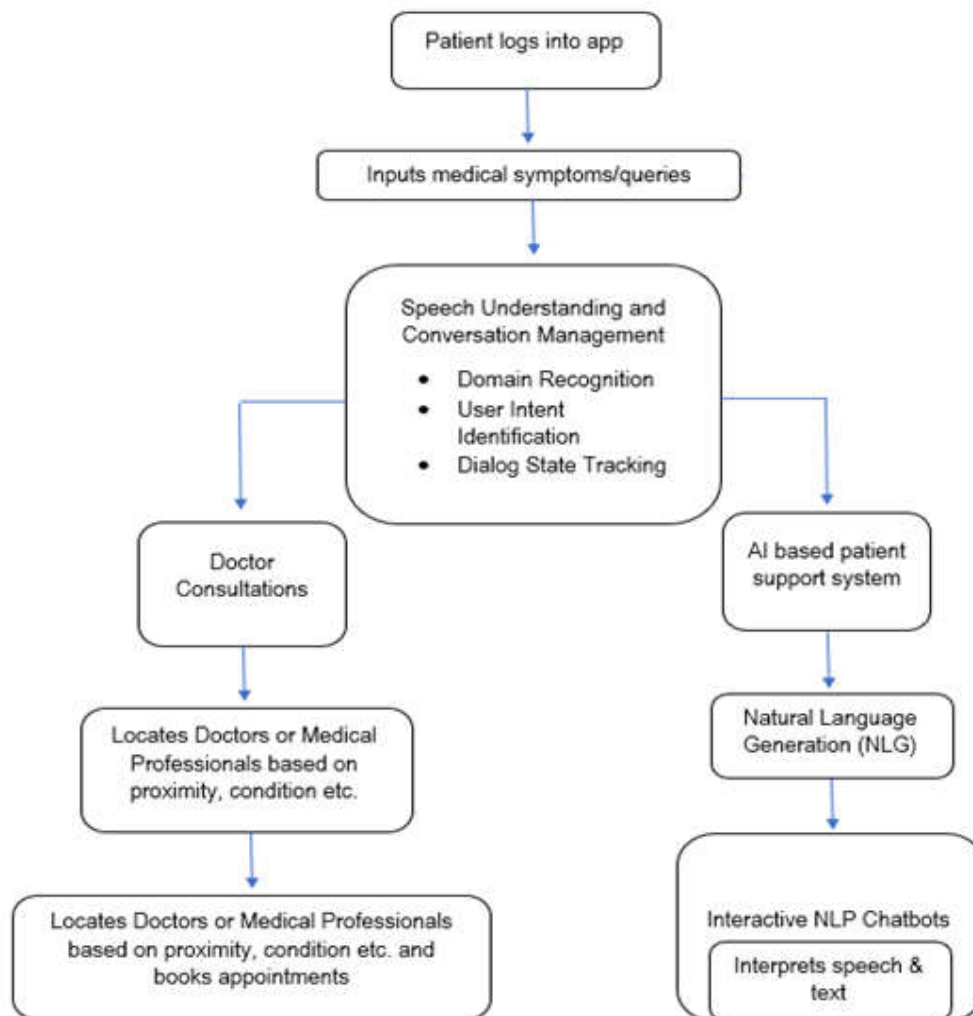


Fig. 4. Healthcare patient emergency service

For an enterprise NLP model to be fully competent and undergo automation, it must train its computational algorithms to incorporate the domain knowledge of an organization's various executives like H.R. employees, Marketing Team, Management Team, Customer Support Teams, etc. This is where the challenge lies as enterprises still haven't been able to access NLP-trained developer models to make modifications on the go. This becomes a cumbersome task in the event of an organizational level restructuring event that is in-line with the business-specific semantics. Moreover, there is a lack of a central unified, standardized model for NLP applications to set some ground rules for any enterprise adopting the Technology and set some global compliance in place, which may hinder the privacy policies of users/clients. This is the need of the hour, as, without a central governing body, NLP applications will soon become like a rudderless boat hovering and drifting on the sea of uncertainty.

## 7. CONCLUSION

While it is clearly evident that Natural Language Programming has evolved from being just a booming technology buzzword in the industry to the most relevant industry in itself and has shown great maturity in these challenging times, our research shows that NLP hasn't just evolved vertically with an array of advanced algorithms and models with more advanced use cases but also grown horizontally across industries and sectors and has shown great promise. We have also seen that the initial adoption rate of the NLP across countries has been low due to the initial higher investment cost. However, it is seen that over the long run, NLP has been found to decrease one's operating cost by reducing redundant work and freeing organization bandwidth for higher productivity. Simply put, NLP provides you with Capex reliability as well as OpEx flexibility. There are thousands of medical documents, notes, and journals present all over the world containing various terminologies that record every step of the healthcare system through clinical testing, training, research, medical treatments, etc., So whenever a global pandemic like Covid-19 strikes, there is an immense need to quickly update the medical protocols and documents accordingly with the feature meticulously treat, document and track the medical patients in an automated manner with much human intervention. This is where NLP solutions have

proven their worth by optimizing and stimulating medical, administrative, and clinical and research protocols providing Quality Assurance, Risk Mitigation, Condition-based Predictive Analytics, and Medical Necessity Review. So, it's true that NLP has a few bottlenecks as of now, but the future looks quite promising.

## CONSENT

It is not applicable.

## ETHICAL APPROVAL

It is not applicable.

## ACKNOWLEDGEMENT

The authors wish to acknowledge Symbiosis Institute of Management Studies for assisting & providing access to resources for this study.

## COMPETING INTERESTS

Authors have declared that no competing interests exist.

## REFERENCES

1. Otter DW, Medina JR, Kalita JK. A survey of the usages of deep learning for natural language processing. *IEEE Transactions on Neural Networks and Learning Systems*; 2020.
2. Gharehchopogh, Farhad Soleimanian, and Zeinab Abbasi Khalifelu. Analysis and evaluation of unstructured data: text mining versus natural language processing. In 2011 5th International Conference on Application of Information and Communication Technologies (AICT), 2011;1-4. IEEE.
3. Top Natural Language Processing Applications in Business. Available: [https://www.accenture.com/\\_acnmedia/PDF-106/Accenture-Unlocking-Value-Unstructured-Data.pdf](https://www.accenture.com/_acnmedia/PDF-106/Accenture-Unlocking-Value-Unstructured-Data.pdf). [Accessed: 10-Aug.-2020].
4. (PDF) NLP challenges in dealing with OCR-ed documents of [Online]. Available: [https://www.researchgate.net/publication/281112670\\_NLP\\_challenges\\_in\\_dealing\\_with\\_OCR-ed\\_documents\\_of\\_derogated\\_quality](https://www.researchgate.net/publication/281112670_NLP_challenges_in_dealing_with_OCR-ed_documents_of_derogated_quality). [Accessed: 10-Aug.-2020].

5. Perera, Rivindu, Parma Nand. Recent advances in natural language generation: A survey and classification of the empirical literature. *Computing and Informatics*. 2017;36(1):1-32.
6. Kang, Ning, Bharat Singh, Zubair Afzal, Erik M. van Mulligen, and Jan A. Kors. Using rule-based natural language processing to improve disease normalization in biomedical text. *Journal of the American Medical Informatics Association*. 2013;20(5):876-881.
7. Kao, Anne, Steve Poteet. Text mining and natural language processing: an introduction for the special issue. *ACM SIGKDD Explorations Newsletter*, 2005;7(1):1-2.
8. Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: a survey. *Ain Shams Eng J*. 2014;5 (4):093–1113.
9. Torfi, Amirina, Rouzbeh A. Shirvani, YaserKeneshloo, Nader Tavvaf, and Edward A. Fox. Natural Language Processing Advancements By Deep Learning: A Survey. arXiv preprint arXiv: 2020;2003.01200.
10. Wu, Yonghui, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, et al. Google's neural machine translation system: Bridging the gap between human and machine translation. arXiv preprint arXiv: 2016;1609.08144.
11. Sit MA, Koylu C, Demir I. Identifying disaster-related tweets and their semantic, spatial and temporal context using deep learning, natural language processing and spatial analysis: a case study of Hurricane Irma. *International Journal of Digital Earth*.
12. Molina, Alejandro, Eric SanJuan, and Juan-Manuel Torres-Moreno. A Turing test to evaluate a complex summarization task. In *International Conference of the Cross-Language Evaluation Forum for European Languages*, pp. 75-80. Springer, Berlin, Heidelberg; 2013.
13. Findings of the 2019 Conference on Machine Translation (WMT19 .... Available: <https://www.aclweb.org/anthology/W19-5301>.
14. Gao, Jianfeng, Michel Galley, Lihong Li, et al. mbox. *Neural Approaches to Conversational A.I. Foundations and Trends® in Information Retrieval*. 2019;13.
15. Minaee, Shervin, NalKalchbrenner, Erik Cambria, NarjesNikzad, MeysamChenaghlu, and Jianfeng Gao. Deep learning-based text classification: A comprehensive review. arXiv preprint arXiv: 2020;2004.03705.
16. Agerri, Rodrigo, German Rigau. Robust multilingual named entity recognition with shallow semi-supervised features. *Artificial Intelligence*, 2016;238: 63-82.
17. Using AI. to unleash the power of unstructured government data. [Online]. Available: <https://www2.deloitte.com/content/dam/Deloitte/lu/Documents/public-sector/lu-ai-unstructured-government-data.pdf>.
18. Wu, Jia-Jhou, and Sue-Ting Chang. Exploring customer sentiment regarding online retail services: A topic-based approach. *Journal of Retailing and Consumer Services*, 2020;55: 102145.
19. Masebe, Norman, MaloseMoseneke, Michelle Burger, and Andries Hennie van Heerden. Digital Disruption in Retail: Management Strategies for South African Shopping Centers. In *International Conference on Applied Human Factors and Ergonomics*. 2020;238-243. Springer, Cham.
20. Wu, Jia-Jhou, Sue-Ting Chang. Exploring customer sentiment regarding online retail services: A topic-based approach. *Journal of Retailing and Consumer Services*, 2020; 55:102145.
21. Linnhoff-Popien C1. *Machine Learning (ML)*, *Digitale Welt*. 2018;2(4)20-21. DOI: 10.1007/s42354-018-0119-x.
22. De Cock C, Milne-Ives M, van Velthoven M, Alturkistani A, Lam C Meinert E. Effectiveness of Conversational Agents (Virtual Assistants) in Health Care: Protocol for a Systematic Review, *JMIR Research Protocols*. 2020;9(3):e16934, DOI: 10.2196/16934.
23. Neves José, Henrique Vicente, Marisa Esteves, FilipaFerraz, AntónioAbelha, José Machado, Joana Machado, JoãoNeves, Jorge Ribeiro, and LúziaSampaio. A deep-big data approach to health care in the AI age. *Mobile Networks and Applications*. 2018;23(4):123-1128.
24. Powles J, Hodson H, Google DeepMind and healthcare in an age of algorithms, *Health and Technology*. 2017;7(4):351-367,

- DOI: 10.1007/s12553-017-0179-1.
25. Sarath priyan U, Ambeth kumar VD, Gunasekaran K. Implementation of E-Health Cloud System Using Timing Enabled Proxy Re-encryption function (RedtPECK). International Journal of MC Square Scientific Research, 2017;9 (2).
26. Demner-Fushman D, Chapman W, McDonald C. What can natural language processing do for clinical decision support, Journal of Biomedical Informatics. 2009;42(5):60-772. DOI: 10.1016/j.jbi.2009.08.007.
27. Mosedale P, Clinical governance, BSAVA Companion, 2011;2:4-7. DOI: 10.22233/20412495.0211.4.]
28. Täckström O, Ganchev K, Das D. Efficient Inference and Structured Learning for Semantic Role Labeling, Transactions of the Association for Computational Linguistics, 2015;3:29-41. DOI: 10.1162/tacl\_a\_00120.
29. MohdZahari M, Moving beyond Protection with PULSE - Using A.I. to make healthcare more affordable and accessible, International Journal of Human and Health Sciences (IJHHS), 2019;16 DOI: 10.31344/ijhhs.v0i0.134.
30. Montenegro J, da Costa C, da Rosa Righi R. Survey of conversational agents in health, Expert Systems with Applications, 2019; 129:56-67, DOI: 10.1016/j.eswa.2019.03.054.

© 2021 Roy et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Peer-review history:*  
*The peer review history for this paper can be accessed here:*  
<http://www.sdiarticle4.com/review-history/69545>